

Optimal Spread in Network Consensus Models

Fern Y. Hunt ^{*}

Applied and Computational Mathematics Division,
National Institute of Standards and Technology, Gaithersburg, Maryland, U.S.A.

Abstract. We consider the problem of identifying a subset of nodes in a network that will enable the fastest spread of information in a decentralized communication environment. For a model of communication based on a random walk on a connected undirected graph, the optimal set over all sets A of the same or smaller cardinality, minimizes $F(A)$ -the sum of the mean first arrival times to the set by random walkers starting at nodes outside of A . This problem was originally posed by Borkar, Nair and Sanketh (2010) who proved that the set function F is supermodular. Unfortunately, the problem as stated is probably NP-complete. In this paper, we introduce an extension of the greedy algorithm that leverages the properties of the underlying graph to produce exact and approximate solutions of pre-defined quality ν . The method requires the evaluation of F for sets of some fixed cardinality m , where m is much smaller than the cardinality of the optimal set. When F has forward elemental curvature κ , we can provide a rough description of the trade-off between solution quality ν and computational effort m in terms of κ .

Keywords: random walk, graph, consensus models, optimal spread, greedoids, supermodular

1 Introduction

The study of information spread (or dually consensus) in complex networks has been the subject of intense research in the past decade for example [18], [19], [13], [20] where the role of distinguished subsets of nodes such as leaders in consensus models and influential spreaders in models of information spread is studied. In addition [20], [13], [1] have developed methods for obtaining optimal spreaders as determined by some measure of subset performance. Secondly a substantial body of related work is concerned with the construction and performance analysis of algorithms for efficient information spread, for example the so-called push/pull algorithms [8], the independent cascade model [1], a random averaging scheme [4] and the GOSSIP model of [5]. In this paper, our focus will be on the first issue: the identification of optimal spreaders in a network. We will use a random walk communication model and an objective function associated with this process. Results of this research are relevant to the design of algorithms for routing in wireless communication systems when location information is not available

^{*} Contribution of the National Institute of Standards and Technology

[19,10], identification of influential individuals in a social network [13] and in sensor placements for efficiently detecting intrusions in computer networks [15].

Given a connected graph $G = (V, E)$ with N vertices V and edges E , information spreads through the network by a process that is dual to the direction of the random walk (see [16]). An optimal spreader in our setting is defined in terms of a set function F defined for a subset $A \subset V$ as the sum of mean first arrival times to A by random walkers that start at nodes outside of A . If A is an effective target set for the random walks (dually an effective spreader) then $F(A)$ is small. Thus the optimal set (subject to a cardinality constraint K) minimizes $F(A)$, subject to $|A| \leq K$,

$$\min_{A \subset V, |A| \leq K} F(A). \quad (1)$$

Recall that a random walker situated at a node $i \in V$, moves to a neighboring node $j \in V$ in a single discrete time step with probability,

$$\text{Prob}\{i \rightarrow j\} = p(i, j) > 0, \quad \text{if } (i, j) \in E, \quad p(i, j) = 0, \quad \text{if } (i, j) \notin E \quad (2)$$

NOTE: In this discussion $p(i, j) = 1/\deg(i)$ where $\deg(i)$ is the degree of node i . However any probabilities for which the resulting Markov chain is ergodic can be used.

The matrix $\mathbb{P} = (p_{ij})_{i,j=1 \dots N}$ is the transition matrix of a Markov chain which in this paper, is assumed to be irreducible and aperiodic ([12]). Starting at node $i \notin A$, a random walker first reaches the set A at a hitting time $T_A = \min\{n > 0 : X_n \in A\}$, where X_n is the node occupied by the walker at time n . Denoting the expected value of this time by $h(i, A) = \mathbb{E}_i[T_A]$, the value of F at A is expressed as

$$F(A) = \left(\sum_{i \notin A} h(i, A) \right). \quad (3)$$

Given A , $F(A)$ can be evaluated by solving a suitable linear equation. Indeed a standard result in Markov chain theory [12] tells us that $h(i, A)$ is the i th component of the vector \mathbf{H} , which solves the linear equation,

$$\mathbf{H} = \mathbf{1} + \mathbb{P}_A \mathbf{H} \quad (4)$$

where $\mathbf{1}$ is a column vector of $N - |A|$ ones and \mathbb{P}_A is the matrix that results from crossing out the rows and columns of \mathbb{P} corresponding to the nodes of A .

Borkar et al [2] showed that for arbitrary subsets $A, B \subseteq V$, $F(A \cap B) + F(A \cup B) \geq F(A) + F(B)$, that is, F is a supermodular function. Thus $-F$ is submodular and our problem is an instance of submodular maximization, a classic problem in combinatorial optimization. In 1987, Nemhauser, Wolsey and Fisher [17] showed for a bounded submodular function that a set constructed by the greedy algorithm has an approximation ratio of $(1 - 1/e)$. More recently, Borgs, Brautbar, Chayes and Lucier [1] and Sviridenko, Vondrak and Ward in [21], showed that approximations of comparable quality could be obtained very efficiently using different methods. To minimize the convergence rate to consensus of a leader-follower network, Clark, Bushnell and Poovendran considered a

supermodular function closely related to ours and showed that the greedy algorithm produces an approximation that is within $(1 - 1/e)$ of optimal. Despite these guarantees we believe the problem in (1) is NP-complete and that a significant improvement in the guaranteed accuracy of the greedy algorithm is unlikely given the problem as presently stated.

In this paper we will discuss a method that obtains an exact or approximate solution to (1) by introducing additional constraints in the problem that are based on properties of the underlying graph. Observing that a vertex cover of the graph with C vertices is an optimal set for $K = C$, sets of cardinality C or less can be assigned a ranking relative to it. Using the rank we define a class of optimal and near optimal sets $L_{\nu,C}$, where ν , the minimum rank of sets in the class, will give a measure of the quality of the approximation. When $m < K$, the class contains the exact solution. Thus to solve the problem for $K < C$, we choose a collection of sets $\mathbf{S} \subset L_{\nu,C}$. Each set in \mathbf{S} has cardinality m where m is the minimum cardinality of sets in $L_{\nu,C}$. The output of this method is the best set that results from a greedy extension of each set in \mathbf{S} , to a set of cardinality K . The method requires the determination of sets of cardinality m each of pre-determined quality ν . When F has elemental curvature κ ([22]), one can answer the questions, given ν , what must m be and conversely given m what ν is achievable?

The plan of the paper is as follows: Section 1 contains a definition and discussion of optimal and near optimal sets ranked relative to a vertex cover of G of cardinality C . We demonstrate how the method is applied to a graph using a collection of sets \mathbf{S} that are subsets of a vertex cover. If every vertex cover contained optimal sets as subsets, it would make sense to use this choice consistently. Unfortunately, that is not the case. This is a reflection of the fact that in general optimality of a set is not preserved by the addition or deletion of elements, otherwise the greedy algorithm would always yield exact solutions. In Section 2, we discuss the properties of the optimal and near optimal sets that are consequences of the supermodularity of F . In particular, we use the properties of the first arrival time to show that F is monotone and as a consequence if $A \in L_{\nu,C}$ then any superset of A is also in $L_{\nu,C}$. We state conditions on $L_{\nu,C}$ that imply that the subsets and supersets of size m in $L_{\nu,C}$ satisfying these conditions form a greedoid (see Section 3 and reference [14]). By choosing \mathbf{S} to be a set of feasible sets of the greedoid, we show empirically in a second example how the method is used to obtain a solution to (1). An important consequence of Section 3 that this example illustrates is that if an exact solution is in the greedoid generated by \mathbf{S} , then our method will find it. We are effectively restricting our search for a solution to feasible optimal and near optimal sets; in particular sets that are reachable by greedy extension of feasible sets. However we do not as yet have sufficient conditions that guarantee the optimal solution is in a particular greedoid. This is a question for future research. In Section 4, the elemental curvature of F , introduced by Wang et al [22] is used to give a rough analysis of the trade-off between the quality of the approximation as measured by ν and the computational effort parameter m .

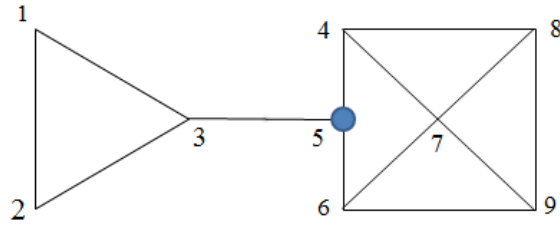


Fig. 1. Graph with $N=9$ vertices, showing optimal set for $K=1$

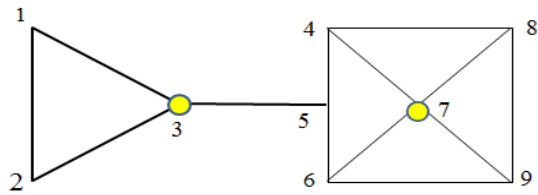


Fig. 2. Optimal set $K=2$ for graph in Figure 1

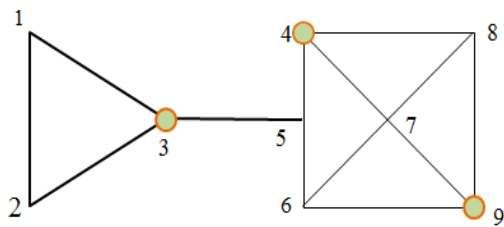


Fig. 3. Optimal set for $K=3$. The set $\{3,6,8\}$ is also optimal by symmetry

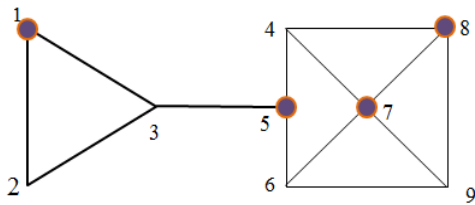


Fig. 4. Optimal Set, $K=4$, for graph in Figure 1

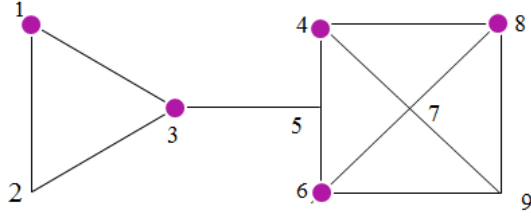


Fig. 5. Optimal Set $K=5$ for graph in Figure 1

$$\begin{array}{l}
 \underline{\mathbf{S}} \\
 \{ \mathbf{3}, \mathbf{7} \} \rightarrow \{ 3, 5, 7 \} \rightarrow \{ 3, 5, 7, 8 \} \rightarrow \{ 1, 3, 5, 7, 8 \} \\
 \{ 1, 7 \} \rightarrow \{ 1, 5, 7 \} \rightarrow \{ \mathbf{1}, \mathbf{5}, \mathbf{7}, 8 \} \rightarrow \{ 1, 3, 5, 7, 8 \} \\
 \{ 3, 8 \} \rightarrow \{ \mathbf{3}, \mathbf{6}, 8 \} \rightarrow \{ 3, 4, 6, 8 \} \rightarrow \{ \mathbf{2}, \mathbf{3}, \mathbf{4}, \mathbf{6}, 8 \} \\
 \{ 3, 6 \} \rightarrow \{ \mathbf{3}, \mathbf{6}, 8 \} \rightarrow \{ 3, 4, 6, 8 \} \rightarrow \{ \mathbf{1}, \mathbf{3}, \mathbf{4}, \mathbf{6}, 8 \}
 \end{array}$$

Fig. 6. Optimal sets for $K=4, 5$ obtained by greedy extension of \mathbf{S}

2 Finding and Approximating Optimal Sets

2.1 Maximal Matches

The optimization problem as posed in equation(1) assumes no advance knowledge about the optimal set or any other possibly related sets. We first consider a process of obtaining optimal sets by using subsets of existing ones. Let A be a vertex cover (not necessarily a minimal one). Since every edge is incident to an element of A , a random walker starting at a vertex i outside of A must hit A at the first step. That is $h(i, A) = 1$. Now equation (4) implies that $h(i, A) \geq 1$ so it follows that A must be an optimal set for its own cardinality. Thus a solution for $C = |A|$ is obtained by constructing a vertex cover. Fortunately a maximal match can be constructed by a simple greedy algorithm and its vertices are a vertex cover with cardinality $C \leq 2 * (\text{cardinality of a minimum vertex cover})$ [7]. Therefore without loss we turn our attention to the solution of problem (1) for $K \leq C$.

2.2 Optimal and Near Optimal Sets

We introduced a measure of the spread effectiveness of sets in section 1, equation (3). It will be convenient to convert this to a rank defined on subsets of V . In particular, suppose there exists a vertex cover with C vertices. We will order

all non-empty subsets $A \subseteq V$ such that $|A| \leq C$ with a ranking function $\rho(A)$ defined as,

$$\rho(A) = \frac{F_{max} - F(A)}{F_{max} - F_{min}} \quad (5)$$

where $F_{max} = \max_{\emptyset \neq A \subseteq V, |A| \leq C} F(A)$, and F_{min} is the corresponding minimum. F_{min} can be calculated by computing F for a maximal match of cardinality C , while F_{max} is the maximal value of F among all one element subsets. We assume that $F_{max} \neq F_{min}$. If this were not the case, $F(A)$ would be have the same value for any non-empty subset A with $|A| \leq C$. Thus any A would be a solution of the problem.

If A is optimal and $|A| = C$ then $\rho(A) = 1$ conversely the the worst performing set has value 0. For a constant ν ($0 < \nu \leq 1$) and C , the non-empty set

$$L_{\nu,C} = \{A : A \subseteq V, |A| \leq C, \rho(A) \geq \nu\} \quad (6)$$

defines a set of optimal and near optimal subsets, with the degree of near optimality depending of course on ν . Let m be the smallest cardinality of sets in $L_{\nu,C}$. Starting with a collection of sets $\mathbf{S} \subset L_{\nu,C}$ of size m , our method is to seek a solution to 1 by greedily augmenting each set until it reaches the desired size K . The offered approximation is the best (has the lowest F value) of these extended sets. We can always find a ν and C such that $L_{\nu,C}$ contains the optimal set of cardinality K but we do not have a proof that the approximation generated by subsets of a vertex cover is optimal. However in section 2, we prove that it is in $L_{\nu,C}$ and therefore has minimum rank ν . We illustrate the method with an example. Figures 1 through 5 show a graph with $N = 9$ vertices along with the vertices of optimal sets for $K = 1$ through 5. To solve the problem for $K = 4$, we note that the class of optimal and near optimal sets based on $C = 8$ and $\nu = .90$ has minimum set size $m = 2$. The set $\mathcal{M} = \{1, 3, 5, 6, 7, 8\}$ is a vertex cover (calculated from the maximal match algorithm). We define \mathbf{S} to be the two element subsets of \mathcal{M} that are in $L_{.90,8}$. The first column of Figure 6 lists these sets and subsequent columns show the results of one element extensions of \mathbf{S} until $K = 5$. Optimal sets are shown in red. In this example the offered approximation is optimal. This is also the case for extensions up to $K = 5$. In this case we see that the method identifies optimal sets that are subsets of \mathcal{M} as well as others that are not e.g. $\{2, 3, 4, 6, 8\}$, underlining the fact the method finds sets that are reachable by greedy extension of subsets of \mathcal{M} . The offered approximation for this method is guaranteed to be in $L_{.90,8}$. This is a consequence of Proposition 1 which is discussed and proved in the next section , Section 3

3 Closure Property of Optimal and Near Optimal Sets

In section 2.2, we demonstrated our method for approximating a solution of problem (1) based on greedy extensions of subsets of a vertex cover that are

optimal or near optimal. Unfortunately a vertex cover can fail to have such subsets other than the vertex cover itself (see an example in [9]). This is the motivation for finding other classes of optimal and near optimal sets that permit the addition and deletion of elements. Our hunch is that greedy extension of such sets will have the largest likelihood of success. The structure we seek is conveniently described in terms of a generalization of the matroid known as a *greedoid* [14,3].

Definition 1 Let \mathbf{E} be a set and let \mathcal{F} be a collection of subsets of \mathbf{E} . The pair $(\mathbf{E}, \mathcal{F})$ is called a greedoid if \mathcal{F} satisfies

- **G1** : $\emptyset \in \mathcal{F}$
- **G2** : For $A \in \mathcal{F}$ non-empty, there exists an $a \in A$ such that $A \setminus \{a\} \in \mathcal{F}$
- **G3** : Given $X, Y \in \mathcal{F}$ with $|X| > |Y|$, there exists an $x \in X \setminus Y$, such that $Y \cup \{x\} \in \mathcal{F}$

A set in \mathcal{F} is called feasible. Note that **G2** implies that a single element can be removed from a feasible set X so that the reduced set is still feasible. By repeating this process the empty set eventually is reached. Conversely starting from the empty set, X can be built up in steps using the **G3** property.

Our first step is to show that $L_{c,K}$ satisfies condition **G3** of the definition for $0 < c \leq 1$, $0 \leq K \leq N$ (Proposition 1). The proof depends on several short lemmas. The first uses an adaptation of an argument in Clark et al [6]

Lemma 1. Let $S \subseteq V$, $u \in V \setminus S$. Then $F(S) \geq F(S \cup \{u\})$.

Proof: Suppose S , a set of nodes is a target set for the random walk. Let $E_{ij}^l(S)$ be the event, $E_{ij}^l(S) = \{X_0 = i \in V, X_l = j \in V \setminus S, X_r \notin S, 0 \leq r \leq l\}$. Thus paths of the random walk start at i and arrive at j without visiting S during the interval $[0, l]$. Also define the event $F_{ij}^l(S, u) = E_{ij}^l(S) \cap \bigcup_{m=0}^l \{X(m) = u\}$ where $u \notin S$. Paths in this event also start at i and arrive at j without visiting S , but must visit the element u at some time during the interval $[0, l]$. Since a path either visits u in the time interval $[0, l]$ or it does not, it follows that:

$$E_{ij}^l(S) = E_{ij}^l(S \cup \{u\}) \cup F_{ij}^l(S, u) \quad (7)$$

We have $E_{ij}^l(S \cup \{u\}) \cap F_{ij}^l(S, u) = \emptyset$. This implies that,

$$\chi(E_{ij}^l(S)) = \chi(E_{ij}^l(S \cup \{u\})) + \chi(F_{ij}^l(S, u)) \quad (8)$$

and therefore:

$$\chi(E_{ij}^l(S)) \geq \chi(E_{ij}^l(S \cup \{u\})) \quad (9)$$

Here $\chi(A)$ is the indicator function of the set A . Recalling that T_S is the hitting time for set S , the following relation comes from taking the expectation of $\chi(E_{ij}^l(S))$ on the left hand side of (9) summing over all $j \in V \setminus S$. Here \mathbb{E} denotes expectation.

$$\text{Prob}\{T_S > l | X_0 = i\} = \mathbb{E} \left(\sum_{j \in V \setminus S} \chi(E_{ij}^l(S)) \right) \quad (10)$$

A similar result is obtained for $T_{S \cup \{u\}}$ from taking the expectation of $\chi(E_{ij}^l(S \cup \{u\}))$ on the right hand side of (9) and summing over $j \in V \setminus S$. Summing once again over all $l \geq 1$ results in the inequality,

$$h(i, S) \geq h(i, S \cup \{u\}) \quad (11)$$

Finally on summing (11) over all i and recalling the definition of F (equation (3)) one obtains the result to be proved. \square

The following result uses that fact that F is supermodular.

Lemma 2. *For $\bar{c} > 0$, let $\mathcal{P} = \{X \subseteq V : F(X) \geq \bar{c} > 0\}$. If $A, B \in \mathcal{P}$ where $|A| = |B|$ and $|A \cap B| = |A| - 1$, then $A \cap B \in \mathcal{P}$.*

Proof: The hypothesis implies the existence of a set X such that $A = X \cup \{a\}$ and $B = X \cup \{b\}$ with $a \neq b, a, b \in V$. The supermodular property of F implies that:

$$F(X \cup \{a\} \cup \{b\}) + F(X) \geq F(X \cup \{a\}) + F(X \cup \{b\}) \quad (12)$$

Rearranging we have,

$$F(X) \geq F(X \cup \{a\}) + [F(X \cup \{b\}) - F(X \cup \{a\} \cup \{b\})] \quad (13)$$

Thus on writing $X = A \cap B$, using the hypothesis on A , and then applying Lemma 1 to the bracketed quantity, we have,

$$F(A \cap B) \geq F(A) \geq \bar{c} \quad (14)$$

\square .

The following lemma is part of a result in [14,3] on paving greedoids.

Lemma 3. *If \mathcal{P} is any class of sets satisfying the conclusion of Lemma 2, $\mathcal{S} = 2^V \setminus \mathcal{P}$ has property **G3**. That is, given any $A, B \in \mathcal{S}$, with $|A| > |B|$, there is an $a \in A \setminus B$ such that $B \cup \{a\} \in \mathcal{S}$.*

Proof: Suppose the conclusion is false. If $|A \setminus B| = 1$, then for $a \in A \setminus B$, $B \cup \{a\} \in \mathcal{P}$. But $A = B \cup \{a\}$. To see this suppose there is some $a' \neq a$ that is not in B . Then $A \setminus B$ contains a' so $|A \setminus B| > 1$ so this is a contradiction. Thus $A = B \cup \{a\}$, but $A \in \mathcal{S}$, and this is also a contradiction. Suppose next that $|A \setminus B| > 1$. Then there exists $a, a' \in A \setminus B$. We have $B \cup \{a\}$ and $B \cup \{a'\} \in \mathcal{P}$. Thus the conclusion of Lemma 2 implies that $B \in \mathcal{P}$, which is a contradiction. \square

Proposition 1. *For $0 < c \leq 1$ and $0 < K \leq N$, let $L_{c,K}$ be the class of sets defined in equation (6). Then $L_{c,K}$ satisfies condition **G3**.*

Proof: If \mathcal{P} is the set defined in Lemma 3 then $L_{c,K} = \mathcal{S}$ for some c . In fact we may set $c = \frac{F_{max} - \bar{c}}{F_{max} - F_{min}}$. If $F_{min} \leq \bar{c} \leq F_{max}$, we have $0 \leq c \leq 1$. $L_{c,K}$ satisfies the conclusion of Lemma 3 and therefore it satisfies property **G3**. \square

The proposition establishes that $L_{c,K}$ satisfies the **G3** property for greedoids. However, **G2** does not hold. For example if the set A has cardinality m where m is the size of the smallest set in $L_{c,K}$ then $A \setminus \{a\}$ cannot be in $L_{c,K}$ for any element $a \in A$. Conversely, let $c_n = \max_{|X| \leq n} \rho(X)$. If $c_m \geq c > c_{m-1}$ then m is the size of the smallest set in $L_{c,K}$. Define G_n to be all sets in $L_{c,K}$ of cardinality n . To create a class of sets with the **G2** property, one constructs subsets of G_m of size $n \leq m$ that are "augmentable" i.e. that satisfy **G3**, while sets G_n for $n > m$ are culled so the remaining sets are supersets of the "augmentable" sets and therefore satisfy **G2**. The greedoid will then consist of selected subsets and supersets of G_m . Conditions for the existence of "augmentable" subsets of G_m and proof of the validity of the resulting greedoid construction can be found in [9]. Rather than repeat the details of these arguments here, we close this section with an example showing the greedoid of a graph (Figure 7) and its use in the solution of (1). The minimum cardinality of a set in the class of optimal and near optimal sets $L_{.85,7}$ is $m = 3$. These sets are used to create the greedoid depicted in Figure 8. Note that **G1-G3** are satisfied. Assume the optimal set for $K = 4$ is unknown. Then our method in this case is to take **S** to be the three element sets in $L_{.85,7}$ that are feasible sets of the greedoid and perform a greedy extension of each set. In figure 8 a line is drawn between a set and its greedy extension. We have also drawn greedy extensions of sets of cardinality $n < m$ as well. The optimal sets are shown in red and so they are in the greedoid. The offered approximations are in fact exact. This example illustrates the fact that if **S** is the class of feasible sets of cardinality m and an optimal set is known to be in the greedoid, then it will be found by our method.

4 Curvature

The greedy extension method discussed in this paper requires computation of the optimal and near optimal sets of cardinality m . Thus it is desirable that m be very much smaller than K . However there is an inherent tradeoff with ν which we take to be a measure of the quality of the approximation. To gain some basic understanding of this tradeoff we will employ the elemental curvature of the rank function (see equation (5)). Elemental curvature was used by Wang, Moran, Wang, and Pan [22] in their treatment of the problem of maximizing a monotone non-decreasing submodular function subject to a matroid constraint. Recall that the objective function F is supermodular, monotone and non-increasing. Together with equation(5) these properties imply that ρ is submodular, monotone and non-decreasing. We note that we can define $F(\emptyset)$ so that ρ can be normalized. Thus we can assume without loss that $\rho(\emptyset) = 0$. The elemental curvature of ρ is defined over $L_{\nu,C}$ in terms of the marginal increase in the rank of a set when a single element is added to it. That is,

$$\rho_j(A) = \rho(A \cup j) - \rho(A). \quad (15)$$

The curvature is then,

$$\kappa = \max\left\{\frac{\rho_i(S \cup j)}{\rho_i(S)} : S \subset L_{\nu,C}, i \neq j, i, j \notin S\right\} \quad (16)$$

Suppose $S \subset T$. For $T \setminus S = \{j_1, \dots, j_r\}$, we have (see equation (2) [22]),

$$\rho(T) - \rho(S) = \sum_{t=1}^r \rho_{j_t}(S \cup \{j_1, \dots, j_{t-1}\}). \quad (17)$$

Therefore ,

$$\rho(T) - \rho(S) \leq \rho_{j_1}(S) + \kappa \rho_{j_2}(S) + \dots + \kappa^{t-1} \rho_{j_r}(S) \quad (18)$$

Suppose $\rho(T) = 1$, for example if T is a vertex cover. Define γ to be $\gamma = \max\{\rho_{j_t}(S) : S \subset T, t = 1 \dots r\}$. We can get a lower bound on the rank of S using equation (18) and the inequality $0 \leq \rho_j(S) \leq \gamma$. First assume γ is known and that $\gamma < 1$. This fact is true is for example when $S \neq \emptyset$. Then,

$$\rho(S) \geq 1 - \gamma \sum_{t=1}^r \kappa^{t-1} \quad (19)$$

Let us now suppose that :

$$(1 - \gamma \sum_{t=1}^r \kappa^{t-1}) \geq \nu, \quad (20)$$

and $|S| \geq m$. If an approximation of quality ν is required, and $r(\nu)$ is the largest value of r such that inequality (20) holds, then the required m must be at least

$$m(\nu) = C - r(\nu) \quad (21)$$

Conversely, given m , the quality of the approximation depends on γ , the largest marginal increase of a set S of size m , κ and $r = C - m$. More precisely, the largest value of ν and thus the guaranteed quality of an approximation obtained by our method , has an upper bound given by the left hand side of (20).

References

1. C. Borgs, M. Brautbar, J. Chayes, B. Lucier, *Maximizing Social Influence in Nearly Optimal Time* Proceedings of the 25th Annual ACM-SIAM Symposium on Discrete Algorithms, 2014, pp. 946-957
2. V.S. Borkar, J. Nair, N. Sanketh, *Manufacturing Consent*, 48th Annual Allerton Conference, Allerton House, UIUC, Illinois, September 2010, pp. 1550-1555
3. A. Björner, G. Ziegler, *Introduction to Greedoids*, in Matroid Applications (ed. N. White), Encyclopedia of Mathematics, Vol. 40, Cambridge University Press, London, UK, 1992, pp.284-357
4. S. Boyd, A. Ghosh, B. Prabhakar, D. Shah, *Randomized Gossip Algorithms*, IEEE Transactions on Information Theory, Vol. 52, No. 6, pp.2508-2530, June 2006

5. K.Censor-Hillel, B. Haeupler, J.Kelner, P. Maymounkow, *Global Computation in a Poorly Connected World:Fast Rumor Spreading with No Dependence on Conductance*, Proceedings of the forty-fourth Annual ACM symposium on the Theory of Computing, 2012, pp.961-970
6. A. Clark, L. Bushnell, R. Poovendran, *Leader Selection for Minimizing Convergence Error in Leader-Follower Systems:A Supermodular Optimization Approach*, 10th International Symposium Modeling and Optimization in Mobile, Ad-Hoc and Wireless Networks (WiOpt), May 2012, pp. 111-115
7. T.H. Cormen, C.E. Leiserson, R.L. Rivest, C. Stein *Introduction to Algorithms*, third edition, MIT Press, 2009
8. G.Giakakoupis, *Tight Bounds for Rumor Spreading with Vertex Expansion*, Proceedings of the 25th ACM-SIAM Symposium on Discrete Algorithms (SODA), 2014, pp.801-815
9. F. Hunt, *The Structure of Optimal and Near Optimal Sets in Consensus Models*,NIST Special Publication 500-303, August 2014, <http://arxiv.org/abs/1408.4364> (online version)
10. A. Jadbabaie, *On geographic routing without location information*,43rd IEEE Conference on Decision and Control, Vol. 5, pp.4764-5769
11. D. Jungnickel, *Graphs, Networks, and Algorithms*, Springer Verlag, New York, Berlin, Heidelberg, Tokyo, 1991
12. J. Kemeny, J. Snell, *Finite Markov Chains*, 2nd edition, Springer-Verlag, New York, Berlin, Heidelberg, Tokyo, 1976
13. D. Kempe, J. Kleinberg, E. Tardos, *Maximizing the Spread of Influence through a Social Network*, Proceedings of the 9th ACM-SIGKIDD International Conference, Washington D.C. 2003, pp.137-146
14. B. Korte, L. Lovasz, R. Schrader, *Greedoids*, Algorithms and Combinatorics Series, Vol. 4, Springer Verlag, Berlin, Germany,1991
15. A. Krause, J. Leskovec, C. Guestrin, J. VanBriesen, C. Faloutsos, *Efficient Sensor Placement Optimization for Securing Water Distribution Networks*, Journal of Water Resource Planning and Management, Vol. 134, No. 6, Nov. 1, 2008
16. R. Lambiotte, R. Sinatra, J.C.Delvenne, T.S. Evans, M. Barahona, V.Lattora, *Interweaving dynamics and structure*, Phys Rev. E **84**, 017102, 2011
17. G.L. Nemhauser, L.A. Wolsey, M.L. Fisher, *An analysis of approximations for maximizing submodular set functions-I*, Mathematical Programming, Vol. **14**, pp.553-574, 2003
18. R. Olfati-Saber, J. Fax, R.M. Murray, *Consensus and cooperation in networked multi agent systems*, Proceedings IEEE, Vol. 95, January 2007, pp. 215-233
19. A.Rao, S.Ratnasamy, C. Papadimitriou, S. Shenker, I. Stoica, *Geographic Routing without Location Information*, Proceedings of the 9th annual international conference on Mobile computing and networking, 2003, pp. 96-108
20. M.Richardson, P.Domingos, *Mining Knowledge Sharing Sites for Viral Marketing*, Eighth International Conference on Knowledge, Discovery and Data Mining, 2002
21. M.Sviridenko, J. Vondrak, J. Ward, *Optimal Approximation for submodular and supermodular with bounded curvature*, accepted SODA15, arXiv:1311.4728v3, December 2014
22. Z.Wang, B.Moran,X.Wang,Q.Pan, *Approximation for maximizing monotone non-decreasing set function with a greedy method*, J.Comb. Optimization (online) DOI 10.1007/s10878-014-9707-3, January 2014

